

Cramér transform of Rademacher series

KRZYSZTOF ZAJKOWSKI ¹

Institute of Mathematics, University of Białystok
Akademicka 2, 15-267 Białystok, Poland
kryza@math.uwb.edu.pl

Abstract

A variational formula for the Cramér transform of series of weighted, independent symmetric Bernoulli random variables (Rademacher series) is given.

2010 Mathematics Subject Classification: 44A15, 60F10

Key words: Rademacher series, Cramér transform, Legendre-Fenchel transform, large deviations

1 Introduction

The *Cramér transform* defines a *rate function* of the *large deviations* for empirical means of a sequence of i.i.d. random variables (see [2]). The literature concerning much more general contexts of the *large deviation principles* is very vast (see for instance monographs [4, 3]). A goal of this paper is only to show some variational formula for the Cramér transform of random variables which are series of weighted, independent symmetric Bernoulli random variables.

The Cramér transform is the *Legendre-Fenchel transform* of the *cumulant generating function* of r.v. We will need the general notion of the Legendre-Fenchel transform in topological spaces (see [5] or [1]). Let X be a real locally convex Hausdorff space and X^* its dual space. By $\langle \cdot, \cdot \rangle$ we denote the canonical pairing between X and X^* . Let $f : X \mapsto \mathbb{R} \cup \{\infty\}$ be a function nonidentically ∞ . By $\mathcal{D}(f)$ we denote the *effective domain* of f , i.e. $\mathcal{D}(f) = \{x \in X : f(x) < \infty\}$. A function $f^* : X^* \mapsto \mathbb{R} \cup \{\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in X} \{\langle x, x^* \rangle - f(x)\} = \sup_{x \in \mathcal{D}(f)} \{\langle x, x^* \rangle - f(x)\} \quad (x^* \in X^*)$$

is called the *Legendre-Fenchel transform* (*convex conjugate*) of f and a function $f^{**} : X \mapsto \mathbb{R} \cup \{\infty\}$ defined by

$$f^{**}(x) = \sup_{x^* \in X^*} \{\langle x, x^* \rangle - f^*(x^*)\} = \sup_{x^* \in \mathcal{D}(f^*)} \{\langle x, x^* \rangle - f^*(x^*)\} \quad (x \in X)$$

¹The author is supported by the Polish National Science Center, Grant no. DEC-2011/01/B/ST1/03838.

is called the *convex biconjugate* of f .

The functions f^* and f^{**} are convex and lower semicontinuous in the weak* and weak topology on X^* and X , respectively. Moreover, the *biconjugate theorem* states that the function $f : X \mapsto \mathbb{R} \cup \{\infty\}$ not identically equal to $+\infty$ is convex and lower semicontinuous if and only if $f = f^{**}$.

Let I be a countable set and $(\epsilon_i)_{i \in I}$ be a Bernoulli sequence, i.e. a sequence of i.i.d. symmetric r.v.'s taking values ± 1 . For $\mathbf{t} = (t_i)_{i \in I} \in \ell^2(I) \equiv \ell^2$ the series

$$X_{\mathbf{t}} := \sum_{i \in I} t_i \epsilon_i$$

converges a.s.. Notice that for $\mathbf{t} \in \ell^1$

$$|X_{\mathbf{t}}| \leq \sum_{i \in I} |t_i| = \|\mathbf{t}\|_1,$$

i.e. $X_{\mathbf{t}}$ is a bounded r.v. and we can define its cumulant generating function on whole \mathbb{R} that is

$$\psi_{\mathbf{t}}(s) = \ln E e^{s X_{\mathbf{t}}}$$

for every $s \in \mathbb{R}$. Because $(\epsilon_i)_{i \in I}$ is i.i.d. Bernoulli sequence then

$$\begin{aligned} \psi_{\mathbf{t}}(s) &= \ln \prod_{i \in I} E e^{s t_i \epsilon_i} \\ &= \ln \prod_{i \in I} \frac{e^{s t_i} + e^{-s t_i}}{2} = \sum_{i \in I} \ln \cosh(s t_i). \end{aligned}$$

Observe that

$$\psi'_{\mathbf{t}}(s) = \sum_{i \in I} t_i \tanh(s t_i).$$

We can not derive an evident form of $\psi_{\mathbf{t}}^*$ by using the *classical Legendre transform* because we can not solve (inverse the derivative $\psi'_{\mathbf{t}}$) the equation

$$\psi'_{\mathbf{t}}(s) = \alpha \tag{1}$$

and find

$$\psi_{\mathbf{t}}^*(\alpha) = \alpha s_{\alpha} - \psi_{\mathbf{t}}(s_{\alpha}),$$

where s_{α} is a solution of the equation (1).

The following theorem shows some variational expression on $\psi_{\mathbf{t}}^*$.

Theorem 1.1. *Let $(\epsilon_i)_{i \in I}$ be a Bernoulli sequence and $\mathbf{t} = (t_i)_{i \in I} \in \ell^1(I)$. The Cramér transform of a variable $X_{\mathbf{t}} = \sum_{i \in I} t_i \epsilon_i$ is given by the following variational formula*

$$\psi_{\mathbf{t}}^*(\alpha) = \min_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \sum_{i \in I} t_i b_i = \alpha}} \psi_1^*(\mathbf{b})$$

for $\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$ and $+\infty$ otherwise, where

$$\psi_1^*(\mathbf{b}) = \frac{1}{2} \sum_{i \in I} \left[(1 + b_i) \ln(1 + b_i) + (1 - b_i) \ln(1 - b_i) \right]$$

is the convex conjugate of a functional $\psi_1 : \ell^1 \mapsto \mathbb{R}$ of the form $\psi_1(\mathbf{t}) = \ln E e^{X\mathbf{t}}$ and $\mathcal{D}(\psi_1^*) \subset \ell_\infty(I)$ denotes its effective domain.

Remark 1.2. Presented in the next section proof techniques are similar, but not the same, to methods used by Ostaszewska and Zajkowski in [6, 7].

2 Proof of Theorem 1.1

We begin with an observation on the absolute value of the cumulant generating function: $|\psi_{\mathbf{t}}(s)| \leq |s| \|\mathbf{t}\|_1$. A parameter \mathbf{t} may be an arbitrary element of ℓ^1 . Formally we can define a function ψ of two variables:

$$\psi(s, \mathbf{t}) = \psi_{\mathbf{t}}(s) = \ln E e^{sX\mathbf{t}} \quad \text{for } (s, \mathbf{t}) \in \mathbb{R} \times \ell^1.$$

Fixing \mathbf{t} or s we write $\psi(s, \mathbf{t}) = \psi_{\mathbf{t}}(s)$ or $\psi(s, \mathbf{t}) = \psi_s(\mathbf{t})$, respectively. First we derive ψ_s^* and next we show how $\psi_{\mathbf{t}}^*$ is expressed by ψ_s^* .

In a standard way one can check the convexity of ψ_s for every $s \in \mathbb{R}$. Let $\mathbf{t}, \mathbf{u} \in \ell^1$ and $\lambda \in (0, 1)$ then

$$\begin{aligned} \psi_s(\lambda \mathbf{t} + (1 - \lambda) \mathbf{u}) &= \ln E e^{s \sum_{i \in I} (\lambda t_i + (1 - \lambda) u_i) \epsilon_i} \\ &= \ln E \left[(e^{s \sum_{i \in I} t_i \epsilon_i})^\lambda (e^{s \sum_{i \in I} u_i \epsilon_i})^{1 - \lambda} \right]. \end{aligned}$$

Using the Hölder inequality for exponents $1/\lambda$ and $1/(1 - \lambda)$ we get

$$E \left[(e^{s \sum_{i \in I} t_i \epsilon_i})^\lambda (e^{s \sum_{i \in I} u_i \epsilon_i})^{1 - \lambda} \right] \leq (E e^{s \sum_{i \in I} t_i \epsilon_i})^\lambda (E e^{s \sum_{i \in I} u_i \epsilon_i})^{1 - \lambda}$$

and, in consequence,

$$\begin{aligned} \psi_s(\lambda \mathbf{t} + (1 - \lambda) \mathbf{u}) &\leq \lambda \ln E e^{s \sum_{i \in I} t_i \epsilon_i} + (1 - \lambda) \ln E e^{s \sum_{i \in I} u_i \epsilon_i} \\ &= \lambda \psi_s(\mathbf{t}) + (1 - \lambda) \psi_s(\mathbf{u}). \end{aligned}$$

Because $\psi_s : \ell^1 \mapsto \mathbb{R}$ and $(\ell^1)^* \simeq \ell_\infty$ then

$$\psi_s^* : \ell_\infty \mapsto \mathbb{R} \cup \{+\infty\}.$$

Let $\mathbf{a} = (a_i)_{i \in I} \in \ell_\infty$. By the definition of the convex conjugate we have

$$\psi_s^*(\mathbf{a}) = \sup_{\mathbf{t} \in \ell^1} \left\{ \langle \mathbf{t}, \mathbf{a} \rangle - \sum_{i \in I} \ln \cosh(st_i) \right\}, \quad (2)$$

where $\langle \mathbf{t}, \mathbf{a} \rangle = \sum_{i \in I} t_i a_i$.

Note that for $s = 0$ we have

$$\psi_0^*(\mathbf{a}) = \begin{cases} 0 & \text{if } \mathbf{a} = \mathbf{0}, \\ +\infty & \text{otherwise.} \end{cases}$$

Assume now that $s \neq 0$. An expression in the curly bracket of (2), denote it by w , is concave and its partial derivatives along vector of basis $e_i = (\delta_{ij})_{j \in I}$ in ℓ^1 (δ_{ij} is the Kronecker delta) equal

$$\frac{\partial}{\partial t_i} w(\mathbf{t}) = \frac{\partial}{\partial t_i} \left(\sum_{i \in I} t_i a_i - \sum_{i \in I} \ln \cosh(st_i) \right) = a_i - s \tanh(st_i).$$

The expression w is a sum of functions with separated variables $(t_i)_{i \in I}$. Concavity of each of these functions implies that the gradient $\nabla w(\mathbf{t}) = (a_i - s \tanh(st_i))_{i \in I}$ belongs to the subgradient $\partial w(\mathbf{t})$ since

$$\forall_{\mathbf{u} \in \ell^1} \quad w(\mathbf{t}) - w(\mathbf{u}) \leq \sum_{i \in I} (t_i - u_i) [a_i - s \tanh(st_i)] = \langle \mathbf{t} - \mathbf{u}, \nabla w(\mathbf{t}) \rangle.$$

The concave function w attained its maximum (global) at the point \mathbf{t} if and only if $\mathbf{0} \in \partial w(\mathbf{t})$. It suffices that

$$\forall_{i \in I} \quad a_i - s \tanh(st_i) = 0.$$

Because $\operatorname{arc} \tanh(x) = \frac{1}{2} \ln \frac{1+x}{1-x}$ for $|x| < 1$ then the partial derivatives equal zero when

$$t_i = \frac{1}{2s} \ln \frac{1 + \frac{a_i}{s}}{1 - \frac{a_i}{s}} \quad \text{for} \quad \left| \frac{a_i}{s} \right| < 1.$$

Substituting the above values of t_i 's into (2) we get

$$\psi_s^*(\mathbf{a}) = \frac{1}{2} \sum_{i \in I} \left[\left(1 + \frac{a_i}{s}\right) \ln \left(1 + \frac{a_i}{s}\right) + \left(1 - \frac{a_i}{s}\right) \ln \left(1 - \frac{a_i}{s}\right) \right] \quad \text{for} \quad \left| \frac{a_i}{s} \right| < 1.$$

Look a bit closely at the effective domain of ψ_s^* that is at the set

$$\mathcal{D}(\psi_s^*) = \left\{ \mathbf{a} \in l_\infty : \psi_s^*(\mathbf{a}) < \infty \right\}.$$

The function $f(x) = (1+x) \ln(1+x) + (1-x) \ln(1-x)$ is even and $f(0) = 0$. Since $\lim_{|x| \rightarrow 1^-} f(x) = 2 \ln 2$ we can extend its domain to the interval $[-1, 1]$. One can check that $(1+x) \ln(1+x) + (1-x) \ln(1-x) \geq x^2$. It follows that

$$\sum_{i \in I} \left[\left(1 + \frac{a_i}{s}\right) \ln \left(1 + \frac{a_i}{s}\right) + \left(1 - \frac{a_i}{s}\right) \ln \left(1 - \frac{a_i}{s}\right) \right] \geq \frac{1}{s^2} \sum_{i \in I} a_i^2$$

and $|a_i| \leq |s|$. Let $\overline{B}_\infty(\mathbf{0}; r)$ denote of the closed ball at the center $\mathbf{0}$ and radius r in the space ℓ_∞ . The properties of f gives that

$$\mathcal{D}(\psi_s^*) \subset \overline{B}_\infty(\mathbf{0}; |s|) \cap \ell^2.$$

Let us note that $\mathcal{D}(\psi_s^*)$ is a symmetric set that is $\mathbf{a} \in \mathcal{D}(\psi_s^*)$ if and only if $-\mathbf{a} \in \mathcal{D}(\psi_s^*)$. Moreover it is symmetric with respect to each coordinates a_i of \mathbf{a} .

Return to the function $\psi_{\mathbf{t}}$. Let us observe that

$$|\psi'_{\mathbf{t}}(s)| = \left| \sum_{i \in I} t_i \tanh(st_i) \right| < \|\mathbf{t}\|_1$$

and $\lim_{s \rightarrow \pm\infty} \psi'_{\mathbf{t}}(s) = \pm \|\mathbf{t}\|_1$. It follows $\mathcal{D}(\psi_{\mathbf{t}}^*) = \psi'_{\mathbf{t}}(\mathbb{R}) = (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$. Because $\psi_{\mathbf{t}}$ is convex and continuous on \mathbb{R} then, by the biconjugate theorem, we get

$$\psi_{\mathbf{t}}(s) = \psi_{\mathbf{t}}^{**}(s) = \sup_{\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)} \{ \alpha s - \psi_{\mathbf{t}}^*(\alpha) \}.$$

On the other hand

$$\psi_{\mathbf{t}}(s) = \psi_s(\mathbf{t}) = \sup_{\mathbf{a} \in \mathcal{D}(\psi_s^*)} \left\{ \langle \mathbf{t}, \mathbf{a} \rangle - \frac{1}{2} \sum_{i \in I} \left[\left(1 + \frac{a_i}{s}\right) \ln \left(1 + \frac{a_i}{s}\right) + \left(1 - \frac{a_i}{s}\right) \ln \left(1 - \frac{a_i}{s}\right) \right] \right\}.$$

If we take $\mathbf{a} = s\mathbf{b}$ then $\psi_s^*(s\mathbf{b}) = \psi_1^*(\mathbf{b})$ with $\mathbf{b} \in \mathcal{D}(\psi_1^*)$. It means that we can rewrite the above variational principle as follows

$$\psi_{\mathbf{t}}(s) = \sup_{\mathbf{b} \in \mathcal{D}(\psi_1^*)} \left\{ s \langle \mathbf{t}, \mathbf{b} \rangle - \frac{1}{2} \sum_{i \in I} \left[(1 + b_i) \ln(1 + b_i) + (1 - b_i) \ln(1 - b_i) \right] \right\}. \quad (3)$$

Take now $\alpha = \langle \mathbf{t}, \mathbf{b} \rangle$. Recall that

$$\sup_{\mathbf{b} \in \overline{B}_\infty(\mathbf{0}; 1)} \langle \mathbf{t}, \mathbf{b} \rangle = \|\mathbf{t}\|_1.$$

We show that every number in $(-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$ is taken by the inner product $\langle \mathbf{t}, \mathbf{b} \rangle$ over the set $\mathcal{D}(\psi_1^*)$. Observe that a vector $\mathbf{b} = \sum_{i \in J} r(\operatorname{sgn} t_i) e_i$, where J is some finite subset of I and $r \in [-1, 1]$, belongs to $\mathcal{D}(\psi_1^*)$ (only finite number of nonzero terms). For this vector we have

$$\langle \mathbf{t}, \mathbf{b} \rangle = r \sum_{i \in J} |t_i|.$$

It follows that the inner product $\langle \mathbf{t}, \mathbf{b} \rangle$ attains over the set $\mathcal{D}(\psi_1^*)$ any number belonging to the interval $(-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$.

For a fixed $\mathbf{t} \in \ell^1$, intersect $\mathcal{D}(\psi_1^*) \subset \ell_\infty$ with a family of hyperplains

$$\left\{ \mathbf{b} \in \ell_\infty : \langle \mathbf{t}, \mathbf{b} \rangle = \alpha \right\}_{\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)}.$$

Now we can divide the supremum of (3) into two parts and get

$$\begin{aligned}
\psi_{\mathbf{t}}(s) &= \sup_{\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)} \sup_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \alpha}} \left\{ s \langle \mathbf{t}, \mathbf{b} \rangle - \psi_1^*(\mathbf{b}) \right\} \\
&= \sup_{\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)} \left\{ s\alpha - \inf_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \alpha}} \psi_1^*(\mathbf{b}) \right\}.
\end{aligned} \tag{4}$$

Define a function

$$\varphi_{\mathbf{t}}(\alpha) = \inf_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \alpha}} \psi_1^*(\mathbf{b}).$$

We prove that in the above definition of function $\varphi_{\mathbf{t}}$ an infimum over the set $\mathcal{D}(\psi_1^*) \cap \{\mathbf{b} \in \ell_{\infty} : \langle \mathbf{t}, \mathbf{b} \rangle = \alpha\}$ is attained and we can replace it by a minimum over this set that is we prove

$$\varphi_{\mathbf{t}}(\alpha) = \min_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \alpha}} \psi_1^*(\mathbf{b}) \tag{5}$$

for $\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$ and $+\infty$ otherwise.

By Banach-Alaoglu theorem the closed (unit) ball $\overline{B}_{\infty}(\mathbf{0}; 1) \subset \ell_{\infty} \simeq (\ell^1)^*$ is weak* compact and for each \mathbf{t} and $\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$ the hyperplane $H_{\mathbf{t}, \alpha} = \{\mathbf{b} \in \ell_{\infty} : \langle \mathbf{t}, \mathbf{b} \rangle = \alpha\}$ is closed in this topology. We have that an intersection $\overline{B}_{\infty}(\mathbf{0}; 1) \cap H_{\mathbf{t}, \alpha}$ is weak* compact. Let ℓ_0 be the space of sequences with finite support. Obviously $\ell_0 \cap \overline{B}_{\infty}(\mathbf{0}; 1) \subset \mathcal{D}(\psi_1^*)$ and $H_{\mathbf{t}, \alpha} \cap \ell_0 \neq \emptyset$. We have

$$\forall_{\mathbf{t} \in \ell^1} \forall_{\alpha \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)} \quad \mathcal{D}(\psi_1^*) \cap H_{\mathbf{t}, \alpha} \supset \overline{B}_{\infty}(\mathbf{0}; 1) \cap H_{\mathbf{t}, \alpha} \cap \ell_0 \neq \emptyset.$$

Recall that the function ψ_1^* is nonnegative and lower semicontinuous in the weak* topology. By Weierstrass Theorem ψ_1^* attains its minimum in the compact set $\overline{B}_{\infty}(\mathbf{0}; 1) \cap H_{\mathbf{t}, \alpha}$. Because an intersection of this set with the effective domain of ψ_1^* is nonempty then it means that a nonnegative infimum is attained at some element in $\mathcal{D}(\psi_1^*)$. It follows that in the definition of $\varphi_{\mathbf{t}}$ we can replace the infimum by minimum and the formula (5) holds.

The formula (4) means that $\psi_{\mathbf{t}}$ is the convex conjugate of $\varphi_{\mathbf{t}}$. To prove an equality $\varphi_{\mathbf{t}} = \psi_{\mathbf{t}}^*$ we should show that $\varphi_{\mathbf{t}}$ is convex and lower semicontinuous.

First we check the convexity of $\varphi_{\mathbf{t}}$. Take $\alpha_1, \alpha_2 \in (-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$. If α_1 or α_2 do not belong to the interval $(-\|\mathbf{t}\|_1, \|\mathbf{t}\|_1)$ then the value of $\varphi_{\mathbf{t}}$ at such α_k equals ∞ and the condition of convexity is trivially satisfied. Let \mathbf{b}_k ($k = 1, 2$) be vectors in $\mathcal{D}(\psi_1^*) \cap H_{\mathbf{t}, \alpha_k}$ such that

$$\varphi_{\mathbf{t}}(\alpha_k) = \min_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \alpha_k}} \psi_1^*(\mathbf{b}) = \psi_1^*(\mathbf{b}_k).$$

Observe that for $\lambda \in (0, 1)$

$$\langle \mathbf{t}, \lambda \mathbf{b}_1 + (1 - \lambda) \mathbf{b}_2 \rangle = \lambda \langle \mathbf{t}, \mathbf{b}_1 \rangle + (1 - \lambda) \langle \mathbf{t}, \mathbf{b}_2 \rangle = \lambda \alpha_1 + (1 - \lambda) \alpha_2,$$

that is $\lambda \mathbf{b}_1 + (1 - \lambda) \mathbf{b}_2 \in H_{\mathbf{t}, \lambda \alpha_1 + (1 - \lambda) \alpha_2}$. The above and convexity of ψ_1^* gives

$$\begin{aligned} \varphi_{\mathbf{t}}(\lambda \alpha_1 + (1 - \lambda) \alpha_2) &\leq \psi_1^*(\lambda \mathbf{b}_1 + (1 - \lambda) \mathbf{b}_2) \\ &\leq \lambda \psi_1^*(\mathbf{b}_1) + (1 - \lambda) \psi_1^*(\mathbf{b}_2) = \lambda \varphi_{\mathbf{t}}(\alpha_1) + (1 - \lambda) \varphi_{\mathbf{t}}(\alpha_2). \end{aligned}$$

Now we prove the lower semicontinuity of $\varphi_{\mathbf{t}}$. Recall that ψ_1^* is convex and lower semicontinuous in the weak* topology on ℓ_{∞} . It means that for any $c \in \mathbb{R}$ the set

$$\{\mathbf{b} \in \ell_{\infty} : \psi_1^*(\mathbf{b}) \leq c\} \quad (6)$$

is weak* closed. Since $\psi_1^* \geq 0$ we can assume that $c \geq 0$. Because the above set is contained in weak* compact unit ball $\overline{B}_{\infty}(\mathbf{0}; 1) \supset \mathcal{D}(\psi_1^*)$ then it is also compact in this topology. Consider a range of the set (6) by the functional $l_{\mathbf{t}} := \langle \mathbf{t}, \cdot \rangle$, i.e.

$$l_{\mathbf{t}}\left(\{\psi_1^*(\mathbf{b}) \leq c\}\right). \quad (7)$$

Since for each $\mathbf{t} \in \ell^1$ the linear functional $l_{\mathbf{t}}$ is continuous on ℓ_{∞} (also in the weak* topology), by the intermediate and extreme value theorems we get that the set (7) is a closed interval. By symmetry of the set (6) and linearity of the functional $l_{\mathbf{t}}$ we get the existence of a real number α such that

$$l_{\mathbf{t}}\left(\{\psi_1^*(\mathbf{b}) \leq c\}\right) = [-\alpha, \alpha].$$

We show that

$$\varphi_{\mathbf{t}}^{-1}((-\infty, c]) = [-\alpha, \alpha].$$

Let $\beta \in \varphi_{\mathbf{t}}^{-1}((-\infty, c])$. Since ψ_1^* is lower semicontinuous, there exists \mathbf{b}_{β} such that

$$c \geq \varphi_{\mathbf{t}}(\beta) = \min_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \beta}} \psi_1^*(\mathbf{b}) = \psi_1^*(\mathbf{b}_{\beta}).$$

That is $\langle \mathbf{t}, \mathbf{b}_{\beta} \rangle = \beta \in [-\alpha, \alpha]$. Conversely, let $\beta \in [-\alpha, \alpha]$. Since $l_{\mathbf{t}} = \langle \mathbf{t}, \cdot \rangle$ is continuous on the connected set $\{\psi_1^*(\mathbf{b}) \leq c\}$, there is $\mathbf{b}'_{\beta} \in \{\psi_1^*(\mathbf{b}) \leq c\}$ such that

$$\langle \mathbf{t}, \mathbf{b}'_{\beta} \rangle = \beta.$$

Note that

$$\varphi_{\mathbf{t}}(\beta) = \min_{\substack{\mathbf{b} \in \mathcal{D}(\psi_1^*) \\ \langle \mathbf{t}, \mathbf{b} \rangle = \beta}} \psi_1^*(\mathbf{b}) \leq \psi_1^*(\mathbf{b}'_{\beta}) \leq c,$$

that is $\beta \in \varphi_{\mathbf{t}}^{-1}((-\infty, c])$.

Because $\varphi_{\mathbf{t}}$ is convex and lower semicontinuous then $\psi_{\mathbf{t}}^* = \varphi_{\mathbf{t}}$, which completes the proof.

Remark 2.1. The result of Theorem 1.1 is similar to those obtained by the *contraction principle* (see for instance [3]) but let us emphasize that we used the space of parameters ℓ^1 to generate the convex conjugate of the investigated function and we did not consider any probability distribution on it.

Remark 2.2. Let us stress that the proof of Theorem 1.1 contains some scheme which allow us to generate, under some assumptions of course, variational formulas on the Cramér transform for another series of random variables.

References

- [1] V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, 4th ed., Springer Monographs in Mathematics, Springer, Dordrecht, 2012.
- [2] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifiques et Industrielles 736 (1938), 5-23. Colloque consacré à la théorie des probabilités, Vol. 3, Hermann, Paris.
- [3] A. Dembo, O. Zeitouni, *Large Deviations Techniques and Applications*. Corrected reprints of the second (1998) edition, Stochastic Modeling and Applied Probability, 38, Springer-Verlag, Berlin, 2010.
- [4] J. D. Deuschel, D. W. Stroock. *Large Deviations*. Pure and Applied Mathematics, 137, Academic Press, Inc., Boston, 1989.
- [5] I. Ekeland, R. Témam, *Convex Analysis and Variational Problems*, Translated from French. Corrected reprint of the 1976 English edition. Classics in Applied Mathematics, 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.
- [6] U. Ostaszewska, K. Zajkowski, *Cramér transform and \mathbf{t} -entropy*, Positivity 18 (2014), no. 2, 347-358.
- [7] K. Zajkowski, *Convex conjugates of analytic functions of logarithmically convex functional*, J. Convex Anal. 20 (2013), no. 1, 243-252.